Sudharshan S. Vazhkudai

Oak Ridge National Laboratory
vazhkudaiss@ornl.gov, http://users.nccs.gov/~vazhkuda, Ph : 865-258-8986

1. CURRENT POSITION

Group Leader, Technology Integration, Oak Ridge National Laboratory (ORNL)

- *Occupation: R&D Management*

2. PROFILE

- ***Experience***: 15 years of experience in the US Department Of Energy's (DOE) national lab system, working in the supercomputing and extreme-scale data center programs. Leading projects in several areas such as high-performance computing (HPC), file and storage systems, non-volatile memory (NVRAM), large-scale data management, analytics appliance, multicores, and distributed computing.
- ***Leadership***: Lead a group of 27 research staff members (bachelors, masters, doctorate-level with 5-30 years R&D experience) to build and deploy robust storage system software/hardware solutions, system software infrastructure, and advanced data management for the nation's premier supercomputing and data center, the Oak Ridge Leadership Computing Facility (OLCF). OLCF comprises of some of the world's fastest systems, and provides billions of computing core hours to a national scientific user base from academia, government agencies, and industry, to perform breakthrough research in climate, materials, alternative energy sources and energy storage, chemistry, nuclear physics, astrophysics, quantum mechanics, and the gamut of scientific inquiry.
- ***Technology Innovation and Creativity***:
  - Lead the TechInt group in the development of the following components for OLCF:
    - ✓ *Systems Software*: Develop systems software for the world's No. 2 supercomputer, Titan (27 petaflop heterogeneous CPU/GPU machine with 560,640 cores), in areas such as the I/O subsystem, NVRAM, resilience, system architecture, data management, and networking.
    - ✓ *Storage Systems*: Deployed the extreme-scale Lustre-based parallel file system (PFS), Spider, which is one of the world's fastest storage systems (in the top 3, with 1 TB/s I/O throughput and 40 PB disk storage capacity). Developing the large-scale disk-cache/tape-based archival storage system software, HPSS (HPSS at OLCF stores 50 PB, with 60+ million files).
    - ✓ *Future Systems*: Technology development for future supercomputers (e.g., O(100) petaflop machine in 2018 and an exaflop machine in 2023), and storage systems (e.g., advanced data hierarchy, comprising of byte-addressable NVRAM, SSD-based burst buffers, object-based PFS, and archival storage).
  - Led the development of a distributed data management solution for the US DOE's Spallation Neutron Source (SNS), a billion dollar national infrastructure, producing O(100 TB) of data.
  - 60+ research papers in highly selective peer-reviewed conferences/journals, with around 1433 citations and an H-index of 16.
- ***Program Development***: Initiated multi-institutional projects worth several million dollars, funded by federal agencies such as the National Science Foundation (NSF), the Department of Energy, and the National Institutes of Health (NIH).
- ***Budget Management***: Manage a multi-million dollar annual budget (~ $6M) and a large-scale storage acquisition budget every four years (> $12M).
- ***Education***: A doctorate in computer science with specialization in distributed storage.
- ***Results-oriented***: Ability to take research ideas to products. Experience delivering innovative products under tight time and budget constraints. Software products deployed on national infrastructure.
- ***Mentoring & Team Building***: Excellent people management and communication skills. Ability to attract and retain talent. Successful in building dynamic teams to achieve both short/long-term goals. An "*Outstanding Mentor Award*" from the Oak Ridge Institute for Science and Education.
- Experience dealing with sensitive vendor and customer relationships.

3. EDUCATIONAL EXPERIENCE

Doctor of Philosophy in Computer Science, May 2003

Institution: University of Mississippi/Argonne National Laboratory (ANL)
Dissertation: Bulk Data Transfer Forecasts and the Implications to Grid Scheduling
Highlight: Wallace Givens fellowship from ANL; Thesis work on the Globus grid toolkit.

Master of Science in Computer Science, December 1998
Institution: University of Mississippi
Thesis: Performance Oriented Distributed OS-Evolutionary Steps towards a Distributed Linux

Bachelor of Engineering in Computer Science, June 1996
Institution: Karnatak University, India

4. WORK EXPERIENCE

OAK RIDGE NATIONAL LABORATORY (US DOE Lab) **Oak Ridge, TN**

2012 – present *Group Leader*, Technology Integration, National Center of Computational Sciences (NCCS)

2003 – 2012 *Research Scientist*, Computer Science and Mathematics Division (CSMD)

THE UNIVERSITY OF TENNESSEE **Knoxville, TN**

2010 – present *Joint Faculty*, Joint Institute of Computational Sciences (JICS)

2005 – 2006 *Adjunct Assistant Professor*, Computer Science

ARGONNE NATIONAL LABORATORY (US DOE Lab) **Argonne, IL**

2000 – 2002 *Givens Fellow/Doctoral Fellowship*, Math and Computer Science Division

THE UNIVERSITY OF MISSISSIPPI **Oxford, MS**

1997 – 2000 *Instructor/Research Assistant*, Department of Computer Science

WIPRO INFOTECH **Bangalore, India**

1996 – 1997 *Design Engineer, R&D*

5. R&D MANAGEMENT

- **Group Leader, Technology Integration, Oak Ridge National Laboratory**

Mission: The Technology Integration (TechInt http://techint.nccs.gov) group is charged with delivering new technologies in supercomputing and storage systems for the Oak Ridge Leadership Computing Facility (OLCF). The group's technology scope includes parallel file systems, non-volatile memory, architecture, archival storage, data management, and networking. The group is responsible for delivering system software solutions for the Titan supercomputer (world's No. 2 machine with 27 petaflops, 560,640 cores, 18,688 compute nodes, and 710 TB DRAM), deploying one of the world's fastest Lustre PFS from the block storage up, building the HPSS archival storage software, and technology development for future O(100 petaflop) acquisitions. We also research and evaluate emerging technologies in the aforementioned areas, and provide the systems programming to seamlessly integrate technologies and tools into the infrastructure as they are adopted.

Composition: 27 research staff, with 12 Ph.D., 11 M.S., and 4 B.S. degrees, and experience ranging from 5-30+ years in systems R&D.

Role: Involves hands-on architecting of technology solutions, program and people management, and maintaining vendor/customer relationships.

- Set a vision for the group, identify R&D directions to pursue by ascertaining gaps in the systems software/hardware stack, conceive project ideas, and conceptualize the basic high-level system design.
- Develop project proposals to secure grants from funding agencies where appropriate.
- Build teams for projects, prioritize and assign tasks to staff, monitor progress of deliverables, and mentor staff on technical roadblocks.
- Architect solutions along with the teams, and participate closely on design deep-dives and reviews.
- Responsible for the successful deployment of products, and engage in outreach to the R&D community/users to improve usage.
- Publish high-quality peer-reviewed papers where appropriate, and contribute to program reviews.
- Employee performance management, appraisals, hiring, retention, and compensation planning.
- Responsible for relationship management with vendors (e.g., large-scale storage acquisition involves disks, networking, and computer hardware companies), customers (diverse user base that uses our tools and systems), and funding agencies.

Growth: In the years as the group lead, I have expanded the group to include more areas other than parallel file systems and archival storage. This includes a significant R&D investment into NVRAM, system

architecture, and data management, which involved successfully lobbying for a need to grow in these areas to senior management as well as writing grant proposals to funding agencies.

Impact of group's R&D:

- Software products have been deployed in the OLCF environment (Titan, Spider, and other clusters), other DOE labs, industry and several sites worldwide, and being used by thousands of users.
- Our tools have resulted in the efficient use of supercomputer time, which is a precious, rigorously peer-reviewed commodity.
- Software contributions have been adopted into the mainstream codebase by several vendors such as IBM and Intel.
- Our processes on storage system acquisition have been adopted by other labs as best practices.
- Our work has been published in high-quality venues, and cited in industry blogs on storage systems.

Thrust Areas: Lead projects in several thrust areas, towards a robust supercomputing infrastructure.

A. File and Storage Systems:

- *Storage system deployment*:
  - ✓ Procurement and deployment of one of the world's fastest, extreme-scale, HPC center-wide storage system, *Spider*, from the block storage up, and its integration to the Titan environment. Config: 1 TB/s throughput, 40 PB capacity, 36 storage system units (SSUs) with 20,160 2 TB near-line SAS disks, organized as 2016 object storage targets (OST) and 288 object storage servers (OSS); integration with Titan supercomputer's 18,688 compute nodes via a scalable I/O routing network infrastructure that comprises of strategically placed 432 I/O routers on the 3D-Torus interconnect fabric, which route the I/O traffic from the compute nodes to the backend storage system. Spider serves the Titan machine as well as other clusters.
  - ✓ Experience in RFP development/evaluation, I/O workload benchmarking, end-to-end performance tuning of the various storage system layers and performance QA, simulation studies to ensure fairsharing of the I/O bandwidth for all compute nodes.
- *Parallel File System*: Contributions to the Lustre open-source PFS in areas such as dynamic striping, balanced I/O placement based on OST load, checksums, temporal replication, novel data recovery by reconstructing from source in case of unrecoverable failure, use of GPUs for cost-effective distributed RAID, and highly parallel tools for file system operations (e.g., Spider PFS with ~ 600 million files). These tools are making a significant impact on both OLCF users as well as the community.
- *Scalable I/O Monitoring*: A large-scale I/O monitoring infrastructure by tracking the workload on the backend storage controllers of the Spider storage, and exposing the throughput and load statistics via databases for higher-level services such as I/O-aware tools.
- *I/O-aware Scheduling*: Automatic extraction of application I/O signatures from noisy storage server-side logs using statistics, use of I/O signatures along with the job scheduler to interleave applications to avoid I/O contention, use of storage controller load to select file system partition for a given job, and a coordinated scheduling mechanism to coincide data staging, computation and data offloading in an attempt to view the PFS as a cache.
- *Modeling Storage System Reliability and Provisioning*: An end-to-end simulator to model the Spider storage system reliability and availability. The framework accommodates the various components and subsystems, their interconnections, failure patterns and propagation, and performs dependency analysis to capture a wide-range of failure cases, and predict data unavailability. The framework is also used to provision spares, based on component failure rates and repair times, as well as analyze storage system capacity/bandwidth provisioning.
- *Archival Storage*:
  - ✓ Development of the HPSS disk-cache/tape storage system software as part of a consortium that comprises of multiple DOE labs and IBM. HPSS is widely used as the archival system by numerous sites. The HPSS deployment at ORNL stores over 50PB and 60+ million files.
  - ✓ Developed a data quality assurance framework to verify the integrity of O(10 million) of files.

B. Non-Volatile Memory (NVRAM): Devising NVRAM solutions for next-generation supercomputers, from three perspectives, namely fault tolerance, memory extension, and active computation, using block-addressable flash on SATA, PCIe, NVMe, and byte-addressable flash on DIMMs.

- *NVRAM-based Burst Buffer Storage*: Collaborating with LLNL and IBM on the design of compute-node local, SSD-based burst buffer for fault tolerance of large-scale applications. The burst buffer will absorb the bulk data at high-speeds while seamlessly draining the data to a PFS.

- *NVMalloc*: A runtime library for applications to use SSDs as a secondary memory partition by explicitly allocating variables on an NVRAM store. Built techniques to map a byte-addressable interface to a block store.
- *Active Flash*: Developed mechanisms to push computation into the flash device controller, to facilitate insitu data analysis, and analyzed performance/energy tradeoffs. Developed several scheduling schemes in FTL to perform on-the-fly analysis. Developed an active storage target framework based on the SCSI T10 OSD-2 specification, and modified exofs in Linux to support active processing.
- *Analyze This*: An analytics workflow-aware storage appliance using an array of active flash devices. Obtained a grant from the DOE's Scientific Data Management program.
  - ✓ Developed a storage system with analysis awareness deeply embedded within every layer: an analysis object abstraction atop the active flash array to tie together the data, and the analytics to be (or was) performed on the data, using the OSD interface and SQLite databases; a workflow scheduler within the storage dispatches analysis jobs to the active flash fabric in a manner that minimizes data movement; a FUSE file system interface, *anFS*, with which users can submit analysis workflow jobs, write data, and monitor the status of the jobs (akin to /proc).
- *Wear leveling*: Conducting a study on the impact on NVRAM lifetime based on multiple dimensions: when the device is used for all of the above workloads, placed either locally on the compute node or centrally on the I/O nodes to be shared by tens of thousands of compute nodes, and using different wear-management schemes in the FTL. Devising techniques to improve wear management based on the insights gained.
- *Processing in byte-addressable NVRAM*: Leading the development of a software shared memory infrastructure atop an array of byte-addressable NVRAM (e.g., PCM), capable of active processing.

C. Data Management:

- *Constellation Scientific Data Workspace*: A graph analytics framework for data analysis and discovery that constructs relationships between several entities such as the millions of data products and their corresponding file system stat metadata, the numerous jobs that produced them over time, the users that ran them, the user-specified metadata and tags, and extracted metadata from the self-describing data itself, so that sophisticated scientific queries can be answered. Overlaid a database layer (in-memory graph database, Cassandra key-value store, and SOLR database index) to capture metadata from the aforementioned data sources, and construct complex relationships. Based on this graph engine, we can answer science queries such as "how did the temperature of the ice sheet vary between six months worth of job runs?" and even predict future data products of interest.
- *DOIs for HPC Data*: Developed a data curation process for long-term storage of HPC data. To this end, developed a Digital Object Identifier workflow as a means to identify and curate important data products. DOI provides a way to cite data products and the associated processes.

D. Computer System Architecture and Resilience:

- *Functional Partitioning of Many-core nodes*: Since supercomputer nodes do not allow timesharing, an end-to-end job (massively parallel simulation on the supercomputer, followed by data analysis on other clusters) incurs several overheads. To mitigate this, we have developed a runtime environment, Functional Partitioning, to partition a many-core node (O(100) cores) such that end-to-end application assist tasks can be scheduled in-situ, alongside the application's simulation component for better end-to-end performance. For example, one can imagine spatially multiplexing data analytics/reduction routines, resilience services, deep-memory data placement, and job monitoring alongside the application's simulation, on the same node, in an attempt to improve the application execution. Tool deployed on Titan, and used by large-scale Climate application jobs.
- *Core Pinning*: Developed an advisory tool to help the multiple MPI processes of a job not compete with each other in sharing the FPUs on the AMD compute node on the Titan supercomputer. This tool is helping tens of thousands of user jobs on a day-to-day basis.
- *Supercomputer Failure Analysis*: Analysis of temporal/spatial failure characteristics of Titan's 560,640 CPU/GPU cores to understand trends in machine failure and MTBF. Based on the insights, devising novel techniques on when to do defensive checkpoint I/O and backfilling jobs.
- *Low-Power Architectures*: Studying low-power architectures such as ARM, ATOM, and their viability in heterogeneous supercomputing.

E. Interconnection Networks: Study the impact of network congestion, I/O routing, and node allocation within a 3D-Torus interconnect fabric on large-scale jobs, spanning Titan's 18,688 nodes (560,640 cores). Validate insights via simulations of different network topologies such as fat tree and dragonfly.

F. **Leadership Metric for jobs:** Given the extreme-scale capability of machines such as Titan, preference is given to large jobs, e.g., one that uses at least 20% of the nodes (defined as a leadership capability metric). However, such a job may not always stress other aspects of the machine such as memory, I/O, and interconnect, which also incur significant acquisition/maintenance costs. Devising a metric for a leadership job that is more inclusive of many of the supercomputer subsystems.
G. **Summit Supercomputer:** Contributed to a larger team in the drafting of the RFP and selection of the next-generation supercomputer at OLCF. This will be an O(100) petaflop machine. Along with the vendor, we are shaping the PFS, SSD burst buffer, and messaging components of the machine.

- **Research Scientist, Computer Science Research, Oak Ridge National Laboratory**

In this role, served as the principal investigator on several storage, and data management projects. Led several multi-institutional projects comprising of several graduate students, post-masters, postdoc, and university faculty researchers. I have been responsible for setting goals, software development, mentoring, monitoring progress, and publishing papers. Served on several masters and doctoral committees of graduate students from North Carolina State University, Virginia Tech, Pennsylvania State University, and Northeastern University. For my mentoring efforts, I was awarded the "*Outstanding Mentor Award*" by the Oak Ridge Institute for Science and Education in 2008. Selected projects are listed below.

A. **FreeLoader Cache using Distributed Storage Scavenging:** Led the design and development of an aggregated, distributed storage infrastructure as a client-side cache by data-intensive applications. The basic idea is the aggregation of space and I/O bandwidth contributions from commodity desktops within a domain to provide a mountable (through FUSE), highly-available, shared cache/scratch space for large, immutable scientific data sets that can be accessed in parallel. Built novel techniques such as asymmetric striping, prefix caching with suffix patching from remote sources, and cost-of-recovery of a dataset to determine its redundancy scheme.
B. **Scientific Data Management for SNS:** Lead architect for a scientific gateway, a distributed computation solution, and data management for O(100 TB) from the Spallation Neutron Source (SNS), a billion dollar infrastructure, catering to a large user community.
C. **Networking:** *Staged Data Transfer*. Expedite the end-user data delivery between HPC centers and end-user locations, thereby alleviating the last-mile problem. Developed decentralized data-offloading and just-in-time staging schemes to move job output and input data by reconciling center purge policies, user delivery and job startup deadlines. Combined point-to-point transfer tools (e.g., GridFTP) and decentralized schemes (e.g., BitTorrent) along with NWS to enable timely data delivery. Using such a scheme, users can exploit advanced networks (100Gb/s) and commodity networks in a seamless fashion to deliver data in time.
D. **Cloud Storage:** Led the development of a FUSE file system atop Azure cloud storage. Developed ways to use the cloud storage as intermediate nodes for data delivery from HPC centers to end-users.

- **Joint Faculty Associate Professor, The University of Tennessee**

As an Associate Professor, I work with students and postdocs on several problems in storage and data management. The National Science Foundation and the National Institutes of Health fund these projects.

A. **Distributed NVRAM-based Checkpoint Storage:** Led a multi-institutional National Science Foundation's High-end Computing File systems and I/O project. Built a distributed, lightweight, intermediate storage for supercomputers, using SSDs, in order to alleviate the I/O bottleneck in checkpointing (periodic bulk writes of program state) and improve fault tolerance. Built novel incremental checkpointing support using checksum comparisons of chunks (de-dup).
B. Technical lead and Co-PI on an NIH effort to build a data repository so that National Cancer Institute (NCI) funded centers can aggregate, annotate, search, index, and download data.

- **Doctoral Thesis at Argonne National Laboratory:** Developed middleware for the orchestration of bulk data transfers in the Globus Data Grid. The middleware comprised of a scalable storage broker and selection heuristics for locating widely replicated data, statistical models and tools to predict the performance of wide-area data transfer times, and techniques for co-allocating transfers.

- **Masters Thesis:** Developed a distributed OS for Linux. Built a high-speed communication protocol for a Linux cluster by short-circuiting the protocol stack; a networked file system, global IPC

mechanism, group communication and remote process execution environment using the communication scheme. Patches for the 2.0 kernel. *Team:* Led several graduate students.

6. SELECTED PUBLICATIONS (full list at http://users.nccs.gov/~vazhkuda/publications.html)

1. D. Tiwari, et. al., "Understanding GPU Errors on Large-scale HPC Systems and the Implications for System Design and Operation", *Proceedings of the 21st IEEE International Symposium on High Performance Computer Architecture (HPCA)*, California, February, 2015.
2. S. Oral, et. al., "Best Practices and Lessons Learned from Deploying and Operating Large-Scale Data-Centric Parallel File Systems", *Proceedings of Supercomputing 2014 (SC14): 27th IEEE/ACM Int'l Conference on High Performance Computing, Networking, Storage and Analysis*, New Orleans, Louisiana, November 2014.
3. D. Tiwari, S. Gupta, S.S. Vazhkudai, "Lazy Checkpointing: Exploiting Temporal Locality in Failures to Mitigate Checkpointing Overheads on Extreme-Scale Systems", *Proceedings of the 44th Annual IEEE/IFIP Int'l Conference on Dependable Systems and Networks (DSN)*, Atlanta, Georgia, June 2014.
4. Y. Liu, R. Gunasekaran, X. Ma, S.S. Vazhkudai, "Automatic Identification of Applications I/O Signatures from Noisy Server-Side Traces", *Proceedings of the 12th USENIX Conference on File and Storage Technologies (FAST)*, Santa Clara, California, February 2014.
5. D. Tiwari, S. Boboila, S.S. Vazhkudai, Y. Kim, X. Ma, P. Desnoyers and Y. Solihin, "Active Flash: Towards Energy-Efficient, In-Situ Data Analytics on Extreme-Scale Machines", *Proceedings of the 11th USENIX Conference on File and Storage Technologies (FAST)*, San Jose, California, February 2013.
6. A. Khasymski, M.M. Rafique, A.R. Butt, S.S. Vazhkudai, D.S. Nikolopoulos, "On the Use of GPUs in Realizing Cost-Effective Distributed RAID," *Proceedings of the IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Washington, D.C., August 2012.
7. C. Wang, S.S. Vazhkudai, X. Ma, F. Meng, Y. Kim, C. Engelmann, "NVMalloc: Exposing an Aggregate SSD Store as a Memory Partition in Extreme-Scale Machines," *Proceedings of the 26th IEEE Int'l Parallel & Distributed Processing Symposium (IPDPS)*, Shanghai, China, May 2012.
8. H. Monti, A.R. Butt, S.S. Vazhkudai, "Timely Result-Data Offloading for Improved HPC Center Scratch Provisioning and Serviceability," *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, Vol. 22, No. 8, pp. 1307-1322, August 2011.
9. R. Prabhakar, S. S. Vazhkudai, Y. Kim, A.R. Butt, M. Li, M. Kandemir, "Provisioning a Multi-Tiered Data Staging Area for Extreme-Scale Machines," *Proceedings of the 31st Int'l Conference on Distributed Computing Systems (ICDCS)*, Minneapolis, MN, June 2011.
10. H. Monti, A.R. Butt, S.S. Vazhkudai, "CATCH: A Cloud-based Adaptive Data Transfer Service for HPC," *Proceedings of the 25th IEEE Int'l Parallel & Distributed Processing Symposium (IPDPS)*, Anchorage, AK, May 2011.
11. M. Li, S. S. Vazhkudai, A.R. Butt, F. Meng, X. Ma, Y. Kim, C. Engelmann, G. Shipman, "Functional Partitioning to Optimize End-to-End Performance on Many-core Architectures," *Proceedings of Supercomputing 2010 (SC10): 23rd IEEE/ACM Int'l Conference on High Performance Computing, Networking, Storage and Analysis*, New Orleans, LA, November 2010.
12. H. Monti, A.R. Butt, S.S. Vazhkudai, "/Scratch as a Cache: Rethinking HPC Center Scratch Storage," *Proceedings of the 23rd ACM Int'l Conference on Supercomputing (ICS)*, Yorktown Heights, NY, June 2009.
13. S.A. Kiswany, M. Ripeanu, S. S. Vazhkudai, A. Gharaibeh, "stdchk: A Checkpoint Storage System for Desktop Grid Computing," *Proceedings of the 28th Int'l Conference on Distributed Computing Systems (ICDCS),* Beijing, China, June 2008.
14. Z. Zhang, C. Wang, S. S. Vazhkudai, X, Ma, G. Pike, F. Mueller, J.W. Cobb, "Optimizing Center Performance through Coordinated Data Staging, Scheduling and Recovery," *Proceedings of Supercomputing 2007 (SC07): Int'l Conference on High Performance Computing, Networking, Storage and Analysis*, Reno, NV, November 2007.
15. S. Vazhkudai, X. Ma, "Recovering Transient Data: Automated On-demand Data Reconstruction and Offloading on Supercomputers," *Operating Systems Review: Special Issue on File and Storage Systems*, Vol. 41, No. 1, pp. 14-18, January 2007.

16. S. Vazhkudai, X. Ma, V. Freeh, J. Strickland, N. Tammineedi, T.A. Simon, S.L. Scott, "Constructing Collaborative Desktop Storage Caches for Large Scientific Datasets," *ACM Transactions on Storage (TOS),* Volume 2, No. 3, pp. 221-254, August 2006.
17. S. Vazhkudai, X. Ma, V. Freeh, J. Strickland, N. Tammineedi, S.L. Scott, "FreeLoader: Scavenging Desktop Storage Resources for Scientific Data," *Proceedings of Supercomputing 2005 (SC'05): Int'l Conference on High Performance Computing, Networking and Storage*, Seattle, WA, November 2005.
18. S. Vazhkudai, J. Schopf, "Using Regression Techniques to Predict Large Data Transfers," *Int'l Journal of High Performance Computing Applications-Special Issue on Grid Computing: Infrastructure and Applications*, Volume 17, No. 3, pp. 249-268, Fall 2003.
19. S. Vazhkudai, J. Schopf, I. Foster, "Predicting the Performance of Wide-Area Data Transfers," *Proceedings of the 16th Int'l Parallel and Distributed Processing Symposium (IPDPS),* Fort Lauderdale, FL, April 2002.
20. S. Vazhkudai, S. Tuecke, I. Foster, "Replica Selection in the Globus Data Grid," *Proceedings of the IEEE Int'l Conference on Cluster Computing and the Grid (CCGRID),* pp. 106-113, Brisbane, Australia, May 2001.

7. PROGRAM DEVELOPMENT ACTIVITIES

I have obtained several external grants from DOE, NSF, and NIH to explore new ideas. Selected list below.

1. L. Ramakrishnan, D. Agarwal, S.S. Vazhkudai, M. Franklin, C. Aragon, "Usable Data Abstractions for Next-Generation Scientific Workflows," ASCR Scientific Data Management Program Announcement LAB 14-1043, $1,200,000, 09/2014-08/2017 (Role: Co-PI).
2. S.S. Vazhkudai, X. Ma, D.K. Panda, "Dynamic Staging Architecture for Accelerating I/O Pipelines," *NSF High-End Computing University Research Activity (HECURA)*, CCF-0937827, $672,223, 04/2010-03/2013 (Role: PI).
3. X. Ma, S.S. Vazhkudai, R. Gunasekaran, G. Shipman, "I/O Coordination to Improve Application Performance Stability on Exa-scale Platforms," *DOE ORNL SEED Funds*, $180,000, FY 2012 (Role: Co-PI).
4. A. Passian, S.S. Vazhkudai, R.H. Farahi, M. Parang, "Physical Sciences-Oncology Center Data Coordination Center (PSOC-DCC)," *NIH-SAIC*, $956,000, FY 2012-2015 (Role: Co-PI).
5. S.S. Vazhkudai, R. Gillen, D.E. Bernholdt, "Evaluating the Role of Cloud Computing for Scientific Discovery," *DOE ORNL Ultrascale LDRD*, $655K, FY 2010-2011 (Role: PI).
6. C. Engelmann, S.S. Vazhkudai, "Soft-Error Resilience for Future-Generation High-Performance Computing Systems," *DOE ORNL LDRD*, $650K, FY 2010-2011 (Role: Co-PI).
7. Y. Jiao, E. Ferragut, S.S. Vazhkudai, M. Hagen, S. Miller, C. Griffin, "Massively Parallel Algorithms for Scalable Exascale Data Analysis," *DOE ORNL LDRD*, $650K, FY 2010-2011 (Role: Co-PI).
8. X. Ma, Y. Zhou, V.W. Freeh, S. Vazhkudai, "Application-adaptive I/O Stack for High End Computing," *NSF HECURA*, CCF-0621470, $266K FY 2007-2009 (Role: Co-PI).
9. S. Vazhkudai, X. Ma, J.W. Cobb, G. Pike, "Storage Virtualization: An Integrated Approach to Machine-Room Storage Management," *DOE ORNL LDRD*, $600K, FY 2007-2008 (Role: PI).
10. C. Engelmann, S.L. Scott, S. Vazhkudai, "Reliability, availability and serviceability (RAS) for terascale computing," *ORNL LDRD*, $576K, FY 2005-2006 (Role: Co-PI).

8. AWARDS

- Outstanding Mentor Award (Feb 2008) - from the Oak Ridge Institute for Science and Education.
- Ph.D. Dissertation Fellowship (2001 - 2002) - Argonne National Laboratory (*Data Grid Research*).
- Wallace Givens Fellowship (2000) - Argonne National Laboratory